This lecture reviews the method for proving concentration inequalities developed by Sourav Chatterjee based on Stein's method of exchangeable pairs. The lecture is based on

`Sourav Chatterjee. Stein's method for concentration inequalities. Probab. Theory Relat. Fields (2007) 138:305-321. DOI 10.1007/s00440-006-0029-y.`

Theorem 42.1. *Let (X, X') be an exchangeable pair on $\mathcal{X}$ (i.e., $d\mathbb{P}(X, X') = d\mathbb{P}(X', X)$. Let $F(x, x') = -F(x, x')$ be antisymmetric, $f(x) = \mathbb{E}(F(x, x')|x)$. Then $\mathbb{E}f(x) = 0$. If further*

$$\Delta(x) = \frac{1}{2}\mathbb{E}\left(\left|\left(f(x) - f(x')\right)F(x, x')\right| \,\middle|\, x\right) \leq Bf(x) + C,$$

then $\mathbb{P}(f(x) \geq t) \leq \exp\left(-\frac{t^2}{2(C+Bt)}\right)$.

Proof.

$$\begin{aligned}
\mathbb{E}\left(h(X)f(X)\right) &\overset{\text{by definition of } f(X)}{=} \mathbb{E}\left(h(X)\cdot\mathbb{E}\left(F(X, X')|X\right)\right) = \mathbb{E}\left(h(X)\cdot F(X, X')\right) \\
&\overset{X, X' \text{ are exchangeable}}{=} \mathbb{E}\left(h(X')\cdot F(X', X)\right) \\
&\overset{F(X, X') \text{ is anti-symmetric}}{=} -\mathbb{E}\left(h(X')\cdot F(X, X')\right) \\
&= \frac{1}{2}\mathbb{E}\left(\left(h(X) - h(X')\right)\cdot F(X, X')\right).
\end{aligned}$$

Take $h(X)$=1, we get $\mathbb{E}f(x) = 0$. Take $h(X) = f(X)$, we get $\mathbb{E}f^2 = \frac{1}{2}\mathbb{E}\left(\left(f(X) - f(X')\right)\cdot F(X, X')\right)$.

We proceed to bound the moment generating function $m(\lambda) = \mathbb{E}\exp(\lambda f(X))$, and use Chebychev's inequality to complete the proof. The derivative of $m(\lambda)$ satisfies,

$$\begin{aligned}
|m'_\lambda(\lambda)| &= \left|\mathbb{E}\left(\underbrace{e^{\lambda f(X)}}_{h(X)}\cdot f(X)\right)\right| \\
&= \left|\frac{1}{2}\mathbb{E}\left(\left(e^{\lambda f(X)} - e^{\lambda f(X')}\right)\cdot F(X, X')\right)\right| \\
&\leq \mathbb{E}\left|\frac{1}{2}\left(\left(e^{\lambda f(X)} - e^{\lambda f(X')}\right)\cdot F(X, X')\right)\right| \\
&\overset{\frac{e^a - e^b}{a-b} = \int_0^1 e^{b+t(a-b)}dt \leq \int_0^1 (te^a + (1-t)e^b)dt = \frac{1}{2}(e^a + e^b)}{\leq} |\lambda|\cdot\mathbb{E}\left(\frac{1}{2}\left(e^{\lambda f(X)} + e^{\lambda f(X')}\right)\cdot\left|\frac{1}{2}\left(f(X) - f(X')\right)\cdot F(X, X')\right|\right) \\
&= |\lambda|\cdot\mathbb{E}\left(e^{\lambda f(X)}\cdot\left|\frac{1}{2}\left(f(X) - f(X')\right)\cdot F(X, X')\right|\right) \\
&= |\lambda|\cdot\mathbb{E}\left(e^{\lambda f(X)}\cdot\underbrace{\mathbb{E}\left(\left|\frac{1}{2}\left(f(X) - f(X')\right)\cdot F(X, X')\right| \,\middle|\, X\right)}_{\Delta(X)}\right) \\
&\leq |\lambda|\cdot\mathbb{E}\left(e^{\lambda f(X)}\cdot\left(B\cdot f(X) + C\right)\right) = |\lambda|\cdot\left(B\cdot m'_\lambda(\lambda) + C\cdot m(\lambda)\right)
\end{aligned}$$

Since $m(\lambda)$ is a convex function in λ, and $m'(0) = 0$, $m'(\lambda)$ always has the same sign as λ. In the interval $0 \le \lambda < 1/B$, the above inequality can be expressed as

$$\begin{aligned}
m'(\lambda) &\le \lambda \cdot (B \cdot m'(\lambda) + C \cdot m(\lambda)) \\
(\log m(\lambda))'_\lambda &\le \frac{\lambda \cdot C}{(1-\lambda B)} \\
\log m(\lambda) &\le \int_0^\lambda \frac{s \cdot C}{1 - s \cdot B} ds \le \frac{C}{1-\lambda \cdot B} \int_0^\lambda s ds = \frac{1}{2} \cdot \frac{C \cdot \lambda^2}{1 - \lambda \cdot B}.
\end{aligned}$$

By Chebyshev's inequality $\mathbb{P}(f(x) \ge t) \le \exp\left(-\lambda t + \frac{1}{2} \cdot \frac{C \cdot \lambda^2}{1-\lambda \cdot B}\right)$. Minimize the inequality over $0 \le \lambda < \frac{1}{B}$, we get $\lambda = \frac{t}{C+Bt}$, and $\mathbb{P}(f(x) \ge t) \le \exp\left(-\frac{t^2}{2\cdot(C+Bt)}\right)$. □

We will use the following three examples to illustrate how to use the above theorem to prove concentration inequalities.

Example 42.2. Let $X_1, \cdots, X_n$ be i.i.d random variables with $\mathbb{E}x_i = \mu_i$, $\text{var}(x_i) = \sigma_i^2$, and $|x_i - \mu_i| \le c_i$. Let $X = \sum_{i=1}^n X_i$, our goal is to bound $|X - \mathbb{E}X|$ probabilistically. To apply the above theorem, take X_i' be an independent copy of X_i for $i = 1, \cdots, n$, $I \sim \text{unif}\{1, \cdots, n\}$ be a random variable uniformly distributed over $1, \cdots, n$, and $X' = \sum_{i \ne I} X_i + X_I$. Define $F(X, X')$, $f(X)$, and $\Delta(X)$ as the following,

$$\begin{aligned}
F(X, X') &\stackrel{\text{def.}}{=} n \cdot (X - X') = n \cdot (X_I - X_I') \\
f(X) &\stackrel{\text{def.}}{=} \mathbb{E}(F(X, X')|X) = \mathbb{E}(n \cdot (X_I - X_I')|X) = \frac{1}{n} \sum_{I=1}^n \mathbb{E}(n \cdot (X_I - X_I')|X) = X - \mathbb{E}X \\
\Delta(X) &\stackrel{\text{def.}}{=} \frac{1}{2} \mathbb{E}\left(|(f(X) - f(X')) \cdot F(X, X')| \,\middle|X\right) \\
&= \frac{n}{2} \cdot \frac{1}{n} \sum_{I=1}^n \mathbb{E}\left((X_I - X_I')^2 \,\middle|X\right) \\
&= \frac{1}{2} \sum_{I=1}^n \left(\mathbb{E}\left(\underbrace{(X_I - \mathbb{E}X_I)^2}_{\le c_i^2} \middle|X\right) + \underbrace{\mathbb{E}\left((X_I' - \mathbb{E}X_I')^2\right)}_{=\sigma_i^2}\right) \\
&\le \frac{1}{2} \sum_{I=1}^n \left(c_i^2 + \sigma_i^2\right).
\end{aligned}$$

It follows that

$$\begin{aligned}
\mathbb{P}\left(\sum X_i - \mathbb{E}\sum X_i \ge t\right) &\le \exp\left(-\frac{t^2}{\sum_i c_i^2 + \sigma_i^2}\right) \\
\mathbb{P}\left(\left(-\sum X_i\right) - \mathbb{E}\left(-\sum X_i\right) \ge t\right) &\le \exp\left(-\frac{t^2}{\sum_i c_i^2 + \sigma_i^2}\right) \\
\mathbb{P}\left(\left|\sum X_i - \mathbb{E}\sum X_i\right| \ge t\right) &\stackrel{\text{union bound}}{\le} 2\exp\left(-\frac{t^2}{\sum_i c_i^2 + \sigma_i^2}\right).
\end{aligned}$$

Example 42.3. Let $(a_{ij})_{i,j=1,\cdots,n}$ be a real matrix where $a_{ij} \in [0,1]$ for $1 \le i,j \le n$, π be a random variable uniformly distributed over the permutations of $1,\cdots,n$. Let $X = \sum_{i=1}^n a_{i,\pi(i)}$, then $\mathbb{E}X = \frac{1}{n}\sum_{i,j} a_{i,j}$, and our goal is to bound $|X - \mathbb{E}X|$ probabilistically. To apply the above theorem, we define exchangeable pairs of permutations in the following way. Given permutation π, pick I, J uniformly and independently from $\{1,\cdots,n\}$, and construct $\pi' = \pi \circ (I,J)$ where (I,J) is a transposition of I and J. The two random variables π and π' are exchangeable. We can define $F(X,X')$, $f(X)$, and $\Delta(X)$ as the following,

$$
\begin{aligned}
F(X,X') &\overset{\text{def.}}{=} \frac{n}{2}(X - X') = \frac{n}{2}\left(\sum_{i=1}^n a_{i,\pi(i)} - \sum_{i=1}^n a_{i,\pi'(i)}\right) \\
f(X) &\overset{\text{def.}}{=} \mathbb{E}\left(F(X,X')|X\right) \\
&= \frac{n}{2}\left(a_{I,\pi(I)} + a_{J,\pi(J)} - a_{I,\pi(J)} - a_{J,\pi(I)}|\pi\right) \\
&= \frac{n}{2}\cdot\frac{1}{n}\sum_I a_{I,\pi(I)} + \frac{n}{2}\cdot\frac{1}{n}\sum_J a_{J,\pi(J)} - \frac{n}{2}\cdot\frac{1}{n^2}\sum_{I,J} a_{I,J} - \frac{n}{2}\cdot\frac{1}{n^2}\sum_{I,J} a_{I,J} \\
&= \sum_i a_{i,\pi(i)} - \frac{1}{n}\sum_{i,j} a_{i,j} \\
&= X - \mathbb{E}X \\
\Delta(X) &\overset{\text{def.}}{=} \frac{1}{2}\cdot\frac{n}{2}\mathbb{E}\left((X - X')^2|\pi\right) \\
&= \frac{n}{4}\mathbb{E}\left(\left(a_{I,\pi(I)} + a_{J,\pi(J)} - a_{I,\pi(J)} - a_{J,\pi(I)}\right)^2|\pi\right) \\
&\le \frac{n}{2}\mathbb{E}\left(a_{I,\pi(I)} + a_{J,\pi(J)} - a_{I,\pi(J)} - a_{J,\pi(I)}|\pi\right) \\
&= X + \mathbb{E}X \\
&= f(X) + 2\mathbb{E}X.
\end{aligned}
$$

Apply the theorem above, and take union bound, we get $\mathbb{P}(|X - \mathbb{E}X| \ge t) \le 2\exp\left(-\frac{t^2}{4\mathbb{E}X + t}\right)$.

Example 42.4. In this example, we consider a concentration behavior of the Curie-Weiss model. Let $\sigma = (\sigma_1,\cdots,\sigma_n) \in \{-1,1\}^n$ be random variables observing the probability distribution

$$G((\sigma_1,\cdots,\sigma_n)) = \frac{1}{Z}\exp\left(\frac{\beta}{n}\sum_{i<j}\sigma_i\sigma_j + \beta\cdot h\sum_{i=1}^n \sigma_i\right).$$

We are interested in the concentration of $m(\sigma) = \frac{1}{n}\sum_i \sigma_i$ around $\tanh(\beta\cdot m(\sigma) + \beta\cdot h)$ where $\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$. Given any σ, we can pick I uniformly and independently from $\{1,\cdots,n\}$, and generate σ'_I according to the conditional distribution of σ_i on $\{\sigma_j : j \ne i\}$ (Gibbs sampling):

$$
\begin{aligned}
\mathbb{P}(\sigma_i' = +1 | \{\sigma_j : j \neq i\}) &= \frac{\exp(\frac{\beta}{n}\sum_{j\neq i}\sigma_j + \beta \cdot h)}{2\cdot(\exp(\frac{\beta}{n}\sum_{j\neq i}\sigma_j + \beta\cdot h) + \exp(-\frac{\beta}{n}\sum_{j\neq i}\sigma_j - \beta\cdot h))} \\
\mathbb{P}(\sigma_i' = -1 | \{\sigma_j : j \neq i\}) &= \frac{\exp(-\frac{\beta}{n}\sum_{j\neq i}\sigma_j - \beta \cdot h)}{2\cdot(\exp(\frac{\beta}{n}\sum_{j\neq i}\sigma_j + \beta\cdot h) + \exp(-\frac{\beta}{n}\sum_{j\neq i}\sigma_j - \beta\cdot h))}.
\end{aligned}
$$

Let $\sigma_j' = \sigma_j$ for $j \neq I$. The two random variables σ and σ' are exchangeable pairs. To apply the above theorem, we define $F(X, X')$, $f(X)$, and $\Delta(X)$ as the following,

$$
\begin{aligned}
F(\sigma, \sigma') &\stackrel{\text{def.}}{=} \sum \sigma_i - \sum \sigma_i' = \sigma_I - \sigma_I' \\
f(\sigma) &\stackrel{\text{def.}}{=} \mathbb{E}\left(F(\sigma,\sigma')|\sigma\right) \\
&= \mathbb{E}\left(\sigma_I - \sigma_I'|\sigma\right) \\
&= \frac{1}{n}\sum_{i=1}^n \mathbb{E}\left(\sigma_i - \sigma_i'|\sigma\right) \text{ , where } \sigma_1', \cdots, \sigma_n' \text{ are all by Gibbs sampling.} \\
&= \frac{1}{n}\sum_{i=1}^n \sigma_i - \frac{1}{n}\sum_{i=1}^n \tanh\left(\frac{\beta}{n}\sum_{j\neq i}\sigma_j + \beta h\right) \\
\Delta(X) &\stackrel{\text{def.}}{=} \frac{1}{2}\mathbb{E}\left(|(f(X) - f(X'))\cdot F(X, X')| \,\middle|\, X\right) \\
&\stackrel{|F(\sigma,\sigma')|\leq 2, |f(\sigma)-f(\sigma')|\leq 2(1+\beta)/n}{\leq} \frac{1}{2}\cdot 2 \cdot \frac{2(1+\beta)}{n}.
\end{aligned}
$$

Thus $\mathbb{P}\left(\left|\frac{1}{n}\sum_i \sigma_i - \frac{1}{n}\sum_i \tanh\left(\frac{\beta}{n}\sum_{j\neq i}\sigma_i + \beta h\right)\right| \geq t\right) \leq 2\exp\left(-\frac{t^2 n}{4(1+\beta)}\right)$. Since $|\tanh(\beta m_i(\sigma) + \beta h) - \tanh(\beta m(\sigma) + \beta h)| \leq \frac{\beta}{n}$ where $m_i(\sigma) = \frac{1}{n}\sum_{j\neq i}\sigma_i$, we have $\mathbb{P}\left(\left|\frac{1}{n}\sum_i \sigma_i - \tanh\left(\beta\cdot m(\sigma) + \beta h\right)\right| \geq \frac{\beta}{n} + \frac{t}{\sqrt{n}}\right) \leq 2\exp\left(-\frac{t^2 n}{4(1+\beta)}\right)$.